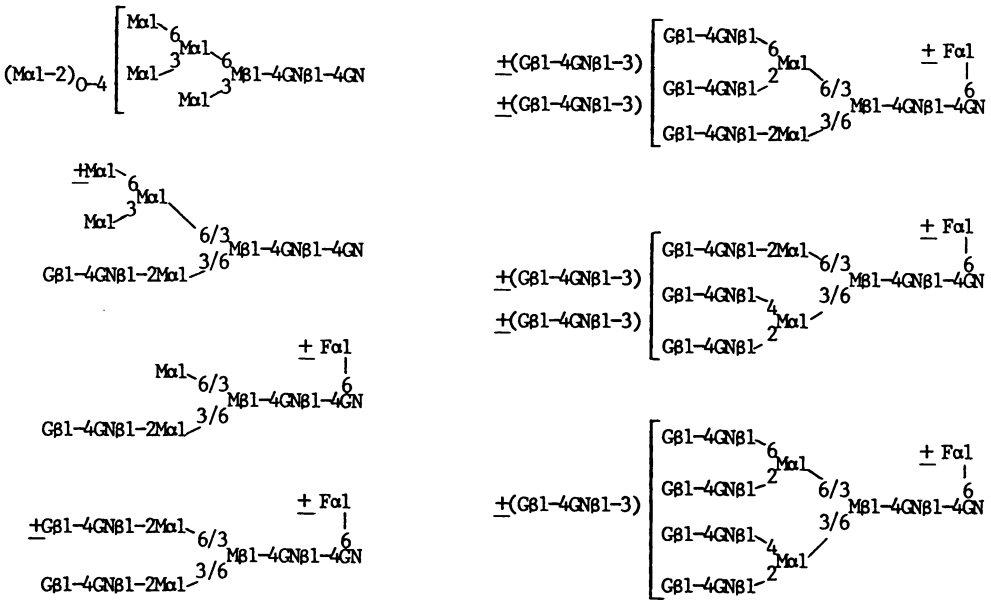
Biochem. J. (1988) 254, 599–603 (Printed in Great Britain)

Carbohydrate structures of the human-immunodeficiency-virus (HIV) recombinant envelope glycoprotein gp120 produced in Chinese-hamster ovary cells

Tsuguo MIZUOCHI,*‡ Michael W. SPELLMAN,† Margot LARKIN,* Julie SOLOMON,* Louisette J. BASA† and Ten FEIZI*§

*Section of Glycoconjugate Research, Medical Research Council Clinical Research Centre, Watford Road, Harrow, Middx. HA1 3UJ, U.K., and †Department of Medicinal and Analytical Chemistry, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080, U.S.A.

The present paper describes the structures of the *N*-linked oligosaccharides of the human-immunodeficiency-virus (HIV) envelope glycoprotein gp120 (cloned from the HTLV-III B isolate and expressed as a secreted fusion protein after transfection of Chinese-hamster ovary cells), which is known to bind with high affinity to human T4-lymphocytes. Oligosaccharides were released from peptide by hydrazinolysis, fractionated by paper electrophoresis, high-performance lectin-affinity chromatography and Bio-Gel P-4 column chromatography, and their structures determined by sequential exoglycosidase digestions in conjunction with methylation analysis. The glycoprotein was found to be unique in its diversity of oligosaccharide structures. These include high-mannose type and hybrid type, as well as four categories of complex-type chains: mono-, bi-, tri- and tetra-antennary, with or without *N*-acetyl-lactosamine repeats, and with or without a core-region fucose residue. Among the sialidase-treated oligosaccharides, no less than 29 structures were identified as follows:



where G is galactose, GN is *N*-acetylglucosamine, M is mannose, F is fucose, and '±' means that residues are present in a proportion of chains. The actual number of oligosaccharide structures is much greater, since before desialylation there was evidence that, among the hybrid and complex-type chains, all but 6% contained sialic acid at the C-3 position of terminal galactose residues, and partially sialylated forms of the bi- and multi-antennary chains were present. Detailed evidence for the proposed oligosaccharide sequences will be published as a supplementary paper [T. Mizuochi, M. W. Spellman, M. Larkin, J. Solomon, L. J. Basa & T. Feizi (1988) Biomed. Chromatogr., in the press].

Abbreviations used: HIV, human immunodeficiency virus; CHO, Chinese-hamster ovary.

‡ On leave of absence from the Division of Biomedical Polymer Science, Institute of Comprehensive Medical Science, Fujita-Gakuen Health University School of Medicine, Toyoake Aichi 470-11, Japan.

§ To whom correspondence and reprint requests should be addressed.

INTRODUCTION

The envelope glycoprotein, gp120, of the human immunodeficiency virus (HIV) is a highly glycosylated molecule [1] with a key role in viral attachment and initiation of infection [2,3] through interaction with the CD4 glycoprotein of T-lymphocytes [4,5]. The oligosaccharides of gp120 are likely to be prominent structures on the virus surface and are candidate attachment and addressing factors in the host. Several observations are consistent with this idea. First, extensive deglycosylation of purified gp120 resulted in impaired binding to T-lymphoblastoid cells [6]. Secondly, cultivation of the virus in the presence of inhibitors of early oligosaccharide processing (glucosidase inhibitors) interfered with HIV infectivity [7]. Thirdly, site-specific mutagenesis at a potential *N*-linked glycosylation site within a conserved domain of gp120 gave rise to viral particles that were non-infectious [8]. Taken together, these observations suggest that there may be a requirement for proper glycosylation both of viral and host-cell glycoproteins for HIV infection, although the precise roles of oligosaccharides are yet to be determined. As part of a programme of investigations into oligosaccharide recognition at various stages of HIV infection, we have determined the structures of *N*-linked oligosaccharides released by hydrazinolysis from a glycosylated soluble recombinant form of gp120 (rgp120) which is produced in Chinese-hamster ovary (CHO) cells [9] and binds with high affinity to human T4-lymphocytes [3]. We report here the oligosaccharide structures identified on rgp120 by enzymic micro-sequencing and methylation analysis.

MATERIALS AND METHODS

Chemicals and enzymes

NaB^3H_4 (15 Ci/mmol) was purchased from New England Nuclear Research Products, Stevenage, Herts., U.K. Details of oligosaccharide standards and the glycosidases [sialidase from *Arthrobacter ureafaciens*; β-galactosidase, β-*N*-acetylhexosaminidase and α-mannosidase purified from jack bean (*Canavalia ensiformis*); β-galactosidase and β-*N*-acetylhexosaminidase from *Diplococcus pneumoniae*; α-mannosidase from *Aspergillus saitoi*; β-mannosidase from *Achatina fulica*; and α-L-fucosidase from bovine epididymis] are given in a supplementary publication [10].

Analytical methods

Paper chromatography, high-voltage paper electrophoresis, Bio-Gel P-4 (extrafine) column chromatography and enzymic micro-sequencing procedures were as described previously [11–13]. Details of the methylation analysis are given in a supplementary publication [10].

Oligosaccharides from rgp120

The rgp120 was purified by immuno-affinity chromatography [9] and gave a single broad band on sodium dodecyl sulphate electrophoresis (results not shown). Since in preliminary analysis of the monosaccharide composition of rgp120 *N*-acetylgalactosamine was not detected, structural studies of the *N*-linked oligosaccharides were performed. A sample (1 mg) of rgp120 was subjected to hydrazinolysis and the oligosaccharides were separated from peptides by paper chromatography [14]. A quarter of the oligosaccharide fraction was reduced in the presence of 3 μmol of NaB^3H_4, followed by 54 μmol of $NaBH_4$ (yield: 6.27×10^7 c.p.m.) and the remainder was reduced with 162 μmol of $NaBH_4$ for methylation analysis.

Affinity chromatography

Affinity chromatography of oligosaccharides (approx. 6×10^5 c.p.m.) using a β-galactoside-specific *Ricinus communis* (castor-bean) agglutinin (RCA 120)-affinity h.p.l.c. column (4.6 mm × 150 mm), from Honen Oil Co. (Tokyo, Japan), was performed as described previously [15], using a Spectra-Physics SP-8700 system, except that phosphate-buffered saline, pH 7.4 (Dulbecco A; Flow Laboratories, Irvine, Ayrshire, Scotland, U.K.), was used instead of 0.01 M-sodium phosphate buffer, pH 7.2, containing 0.15 M-NaCl. After repeated chromatography, pooled oligosaccharide fractions were desalted with Bio-Rad AG-50 (H^+ form) and Bio-Rad AG-3 (OH^- form) resins.

RESULTS AND DISCUSSION

Acidic oligosaccharides found to be sialylated

When the radiolabelled oligosaccharides from rgp120 were subjected to paper electrophoresis at pH 5.4, a neutral fraction, N, and four acidic fractions, A1, A2, A3 and A4, were obtained; these represented 36.9, 25.0, 24.6, 10.9 and 2.6% respectively of the total radiolabelled oligosaccharides. Upon incubation with sialidase, the acidic fractions were completely converted into neutral components (results shown in a supplementary publication [10]), indicating that the acidic property was due to the presence of sialic acid residues.

Oligosaccharide structures identified after desialylation (29 in all)

Before detailed structural studies, radiolabelled neutral oligosaccharides obtained after sialidase treatment of total oligosaccharides were separated into four fractions on the basis of their content of terminal galactose residues by h.p.l.c. using the RCA 120-affinity column: an unretained fraction, I, two retarded fractions, II and III, and a retained fraction, IV, which was eluted in the presence of 0.05 M-lactose (Fig. 1). These represented 32.6, 10.9 and 34.2 and 22.3% respectively of the sialidase-treated total radiolabelled oligosaccharides. Further fractionation of fractions I–IV by Bio-Gel P-4 chromatography (Fig. 2) yielded twelve peaks, **a**–**e** derived from fraction I, **f**–**h** from fraction II, **i** from fraction III, and **j**–**l** from fraction IV. These were pooled as indicated and subjected to enzymic sequencing procedures. Linkage information was deduced from the substrate specificities of the glycosidases used and from methylation analysis of non-labelled oligosaccharides corresponding to fractions I–IV as detailed in ref. [10]. No less than 29 structures were identified among the sialidase-treated oligosaccharides. Their structures and relative amounts (calculated from their radioactivities) are shown in Fig. 3.

Peaks **a**–**e** were identified as high-mannose-type oligosaccharides $Man_5GlcNAc_2$, $Man_6GlcNAc_2$, $Man_7GlcNAc_2$, $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ respectively (Fig. 3). They represented most (approx. 90%) of the neutral oligosaccharides (fraction N) before sialidase treatment. In all, 24 different oligosaccharides

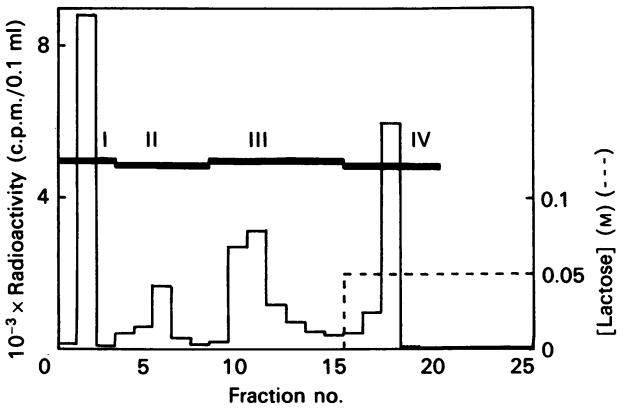


Fig. 1. Separation of the radiolabelled sialidase-treated oligosaccharides on an affinity column of *Ricinus communis* lectin (RCA 120)

The oligosaccharides obtained from rgp120 by hydrazinolysis were radiolabelled, treated with sialidase and chromatographed on an RCA 120-lectin-affinity h.p.l.c. column (4.6 mm × 150 mm; flow rate 0.5 ml/min). The radioactivity in each tube (2 ml/tube) was determined by liquid-scintillation counting. The unretained oligosaccharides (I), the retarded oligosaccharides (II and III) and the oligosaccharides eluted in the presence of 0.05 M-lactose (IV) were pooled as indicated by the bold horizontal bars.

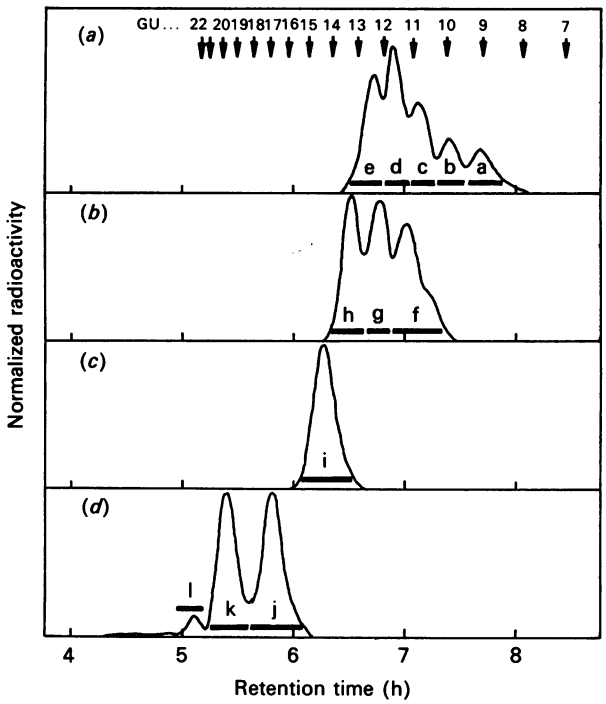


Fig. 2. Chromatography of the radioactive neutral oligosaccharide fractions I–IV on a column of Bio-Gel P-4

The radioactive fractions I (*a*), II (*b*), III (*c*) and IV (*d*) obtained after RCA 120-affinity chromatography were chromatographed on a Bio-Gel P-4 column (1.5 cm × 100 cm; flow rate 0.2 ml/min). The radioactivity in each tube (0.8 ml/tube) was determined by liquid-scintillation counting. Fractions were pooled as indicated by the solid bars (**a**–**l**). Arrows indicate the elution positions of glucose oligomers [12]; values against the arrows indicate the number of glucose units (GU). The radioactivity profiles were 'normalized' so that the heights of the main peak in each panel were about the same.

were identified in peaks **f**–**l**. These were of the hybrid type (oligosaccharides $\mathbf{f_3}$ and $\mathbf{g_1}$), monoantennary complex type ($\mathbf{f_1}$ and $\mathbf{f_2}$), biantennary complex type ($\mathbf{g_2}$, **h** and **i**), triantennary complex type ($\mathbf{j_1}$ $\mathbf{j_2}$, $\mathbf{k_2}$, $\mathbf{k_3}$, $\mathbf{l_2}$ and $\mathbf{l_3}$) and tetra-antennary complex type ($\mathbf{k_1}$ and $\mathbf{l_1}$) and they represented respectively 4.4, 2.4, 38.3, 12.3 and 10.0% of the structures in Fig. 3. *N*-Acetyl-lactosamine repeats were detected in a minority (3%) of the complex-type oligosaccharides; these consisted of single repeats ($\mathbf{k_2}$, $\mathbf{k_3}$ and $\mathbf{l_1}$), or two repeats ($\mathbf{l_2}$ and $\mathbf{l_3}$). Among the complex-type chains, a core-region fucose residue α1-6-linked to reducing-terminal *N*-acetylglucosamine was detected in all of $\mathbf{f_2}$ and **h**, in approx. 90% of **i**–$\mathbf{l_3}$, but not in $\mathbf{f_1}$ and $\mathbf{g_2}$.

Heterogeneity in sialylation patterns detected

Preliminary methylation analysis of total glycopeptides obtained by Pronase digestion of rgp120 revealed C-3-substituted galactose residues in high abundance, non-reducing terminal galactose residues in low abundance, and C-6-substituted galactose residues were not detected. In contrast, among the desialylated oligosaccharides, large amounts of non-reducing terminal galactose residues were detected (details given in ref. [10]), suggesting that the sialic acid residues were linked to the C-3 position of most terminal galactose residues oligosaccharides $\mathbf{f_1}$–$\mathbf{l_3}$. The sialylation patterns among the hybrid-type and complex-type oligosaccharides from rgp120 were investigated, first by comparison of the positions of migration upon paper electrophoresis of the acidic oligosaccharide fractions with those of sialylated mono-, bi-, tri- and tetra-antennary standards, and, secondly, from the elution profiles on the Bio-Gel P-4 column of the neutral oligosaccharide fraction N and those of the desialylated products from A1, A2, A3 and A4. The results (given in ref. [10]) indicated that about 6% of the galactosyl oligosaccharides $\mathbf{f_1}$ to $\mathbf{l_3}$ were non-sialylated, 94% were sialylated and that partially sialylated forms of the bi- and multi-antennary oligosaccharides were present. Thus the total number of oligosaccharide structures before desialylation is predicted to be in the order of 100. No structures that would behave as heterophile antigens immunogenic for man were detected among the oligosaccharides from rgp120.

Comparison with other CHO-cell-derived glycoproteins

When the oligosaccharides of rgp120 reported here are compared with those of recombinant human β- and γ-interferon and recombinant human erythropoietin produced in CHO cells [16–19], there are both structural and quantitative differences: no high-mannose type and hybrid-type oligosaccharides were detected in the latter three proteins, and there were substantial differences in the proportions of the various complex-type chains and the amounts of *N*-acetyl-lactosamine repeats among the four proteins. These observations highlight the marked effect of the protein moiety on the patterns of *N*-linked oligosaccharide processing.

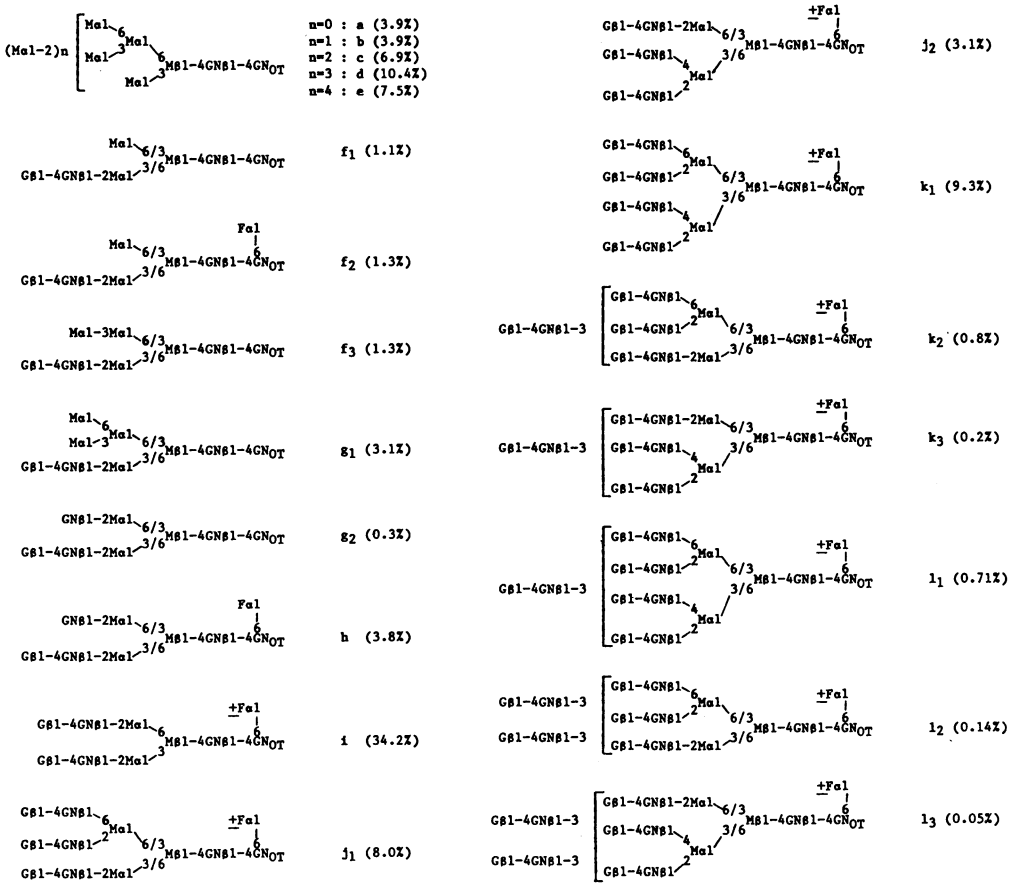


Fig. 3. Structural assignments for oligosaccharides contained in fractions a–l derived from rgp120

Abbreviations are as defined in the Synopsis; '±F' indicates those oligosaccharides of which about 90% are fucosylated. The molar percentages (determined on the basis of radioactivity) refer to the proportions of the total sialidase-treated oligosaccharides **a–l**.

Biological relevance of diversity and abundance of oligosaccharide structures on rgp120

To our knowledge rgp120 is unique among glycoproteins thus far characterized with respect to the diversity of *N*-linked oligosaccharide types that it contains: high-mannose type and hybrid type, in addition to four categories of complex-type chains, mono-, bi-, tri- and tetra-antennary, with and without *N*-acetyllactosamine repeats, and the core-region fucose residue. Preliminary structural studies on natural gp120 derived from H9 lymphoblastoid cells infected with HIV have indicated that the *N*-linked oligosaccharides are similarly diverse (T. Mizuochi, T. Matthews & T. Feizi, unpublished work). Knowledge of this enormous diversity is important in considering the possible involvement of the viral oligosaccharides in interactions with host components at various stages of the virus cycle. The oligosaccharide structures found in such abundance on gp120 are potential ligands for various carbohydrate-binding proteins (endogenous lectins), which are widely distributed in various tissues and secretions [20–22]. Interactions of gp120 with endogenous lectins would result in the coating of the viral envelope with proteins of non-immune origin. It will be important, therefore, to investigate whether such interactions co-operate with, or compromise, the protective immune recognition systems, and whether they enhance the clearance or favour the sequestration of HIV in the infected host.

This work was supported by the Medical Research Council, the Cancer Research Campaign and the Leukaemia Research Fund. We are grateful to Dr. E. F. Hounsell, Dr. R. A. Childs and Dr. R. W. Loveless for helpful discussions and to Mrs. Maureen Moriarty for preparation of the manuscript.

REFERENCES

1. Montagnier, L., Clavel, F., Kurst, B., Chamaret, S., Rey, F., Barré-Sinoussi, F. & Chermann, J. C. (1985) Virology **144**, 283–289
2. McDougal, J. S., Kennedy, M. S., Sligh, J. M., Cort, S. P., Mawle, A. & Nicholson, J. K. A. (1986) Science **231**, 382–385
3. Lasky, L. A., Nakamura, G., Smith, D. H., Fennie, C., Shimasaki, C., Patzer, E., Berman, P., Gregory, T. & Capon, D. J. (1987) Cell (Cambridge, Mass.) **50**, 975–985
4. Dalgleish, A. G., Beverley, P. C. L., Clapham, P. R., Crawford, D. H., Greaves, M. F. & Weiss, R. A. (1984) Nature (London) **312**, 763–766
5. Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.-C. & Montagnier, L. (1984) Nature (London) **312**, 767–768

6. Matthews, T. J., Weinhold, K. J., Lyerly, H. K., Langlois, A. J., Wigzell, H. & Bolognesi, D. P. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 5424–5428
7. Gruters, R. A., Neefjes, J. J., Tersmette, M., de Goede, R. E. Y., Tulp, A., Huisman, H. G., Miedema, F. & Ploegh, H. L. (1987) Nature (London) **330**, 74–77
8. Wiley, R. L., Smith, D. H., Lasky, L. A., Theodore, T. S., Earl, P. L., Moss, B., Capon, D. J. & Martin, M. A. (1988) J. Virol. **62**, 139–147
9. Lasky, L. A., Groopman, J. E., Fennie, C. W., Benz, P. M., Capon, D. J., Dowbenko, D. J., Nakamura, G. R., Nunes, W. M., Renz, M. E. & Berman, P. W. (1986) Science **233**, 209–212
10. Mizuochi, T., Spellman, M. W., Larkin, M., Solomon, J., Basa, L. J. & Feizi, T. (1988) Biomed. Chromatogr., in the press
11. Mizuochi, T., Taniguchi, T., Shimizu, A. & Kobata, A. (1982) J. Immunol. **129**, 2016–2020
12. Yamashita, K., Mizuochi, T. & Kobata, A. (1982) Methods Enzymol. **83**, 105–126
13. Mizuochi, T., Fujii, J., Kisiel, W. & Kobata, A. (1981) J. Biochem. (Tokyo) **90**, 1023–1031
14. Takasaki, S., Mizuochi, T. & Kobata, A. (1982) Methods Enzymol. **83**, 263–268
15. Mizuochi, T., Hamako, J. & Titani, K. (1987) Arch. Biochem. Biophys. **257**, 387–394
16. Conradt, H. S., Egge, H., Katalinic, J.-P., Reiser, W., Siklosi, T. & Schaper, K. (1987) J. Biol. Chem. **262**, 14600–14605
17. Mutsaers, J. H. G. M., Kamerling, J. P., Devos, R., Guisez, Y., Fiers, W. & Vliegenthart, F. G. (1986) Eur. J. Biochem. **156**, 651–654
18. Sasaki, H., Bothner, B., Dell, A. & Fukuda, M. (1987) J. Biol. Chem. **262**, 12059–12076
19. Takeuchi, M., Takasaki, S., Miyazaki, H., Kato, T., Hoshi, S., Kochibe, N. & Kobata, A. (1988) J. Biol. Chem. **263**, 3657–3663
20. Barondes, S. H. (1984) Science **223**, 1259–1264
21. Feizi, T. (1985) Nature (London) **314**, 53–57
22. Drickamer, K. (1987) Kidney Int. **32** (suppl. **23**), 67–80

Received 3 June 1988; accepted 23 June 1988